AD_____

Award Number: DAMD17-03-1-0687

TITLE: Applying Statistical Models to Mammographic Screening Data to Understand Growth and Progression of Ductal Carcinoma in Situ

PRINCIPAL INVESTIGATOR: Dorota Gertig, Ph.D.
Bircan Erbas, Ph.D.
Graham Byrnes, Ph.D.
James Dowty, Ph.D.

CONTRACTING ORGANIZATION: University of Melbourne
Victoria 3010 Australia

REPORT DATE: September 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2004	Annual (18 Aug 03 - 17 Aug 04)

4. TITLE AND SUBTITLE
Applying Statistical Models to Mammographic Screening Data to Understand Growth and Progression of Ductal Carcinoma in Situ

5. FUNDING NUMBERS
DAMD17-03-1-0687

6. AUTHOR(S)
Dorota Gertig, Ph.D.
Bircan Erbas, Ph.D.
Graham Byrnes, Ph.D.
James Dowty, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Melbourne
Victoria 3010 Australia

E-Mail: d.gertig@unimelb.edu.au

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Little is known about the natural history of ductal carcinoma in-situ (DCIS). Estimates from studies of recurrence following surgery suggest about 30% recur as invasive cancer. The aim of this study is to use novel applications of statistical methods to estimate the proportion of DCIS that progress to invasive cancer. We first analysed observed screening data and showed that neither time since previous negative screen or HRT use were associated with size of DCIS. Similar results were found for histological grade. We then developed a computer simulation for mammographic screening data which models progression and detection of Ductal carcinoma in situ. The purpose of the simulation is to infer the distribution of DCIS sizes we would expect in mammography data under different scenarios of tumour initiation, growth, invasion and detection. The simulation therefore provides the user with a test for these competing theories by enabling comparison with actual mammography data. Preliminary results show that low growth rates and low invasion rates provide the best fit to the data. Further work will include the addition of screening round and different mechanisms of invasion to the modeling.

14. SUBJECT TERMS
Mammographic Screening, DCIS, Statistical Models

15. NUMBER OF PAGES
12

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

Little is known about natural history and growth patterns of DCIS. It is estimated that about 30% of DCIS may progress to invasive cancer[1,2,3], based on studies of long-term follow-up of DCIS initially mis-diagnosed as benign, but this cannot be studied directly. Previous models of invasive tumor growth have used the distribution of tumor size (as measured by diameter) at detection and assumed constant growth[4,5]. These models are parameterised by tumor volume; however, these assumptions may not hold for DCIS. The aim of this proposal is to apply novel statistical models to explore the growth patterns of DCIS. Our approach differs from previous models in that we model the DCIS phase (both growth and invasion) and model detection via mammography at discrete intervals. These models may lead to new insights into natural history of DCIS that can be used to predict DCIS growth and may provide important data for clinical management.

BODY

Substantial progress has been made in modeling and simulation of growth of DCIS, in particular for Tasks 2 and 3. Below we describe the progress on each task related to the statement of work. **All data presented are preliminary.**

Task 1:

Validation of measures of DCIS size, comparing reported size on pathology reports with size as measured on mammograms.

Progress on this aspect of the study has been delayed due to difficulty in finding a radiologist associated with the program who has the time commitments to conduct measurements of DCIS size on mammograms in the validation study. Thus we have requested an extension to complete this work. We believe that a radiologist affiliated with the screening program will be able to complete this work within the next 8 months.

Tasks 2 and 3: The progress on Tasks 2 and 3 is described together as they are inherently linked to the modeling and simulation of DCIS growth rates.

Preparation of data set:
We received de-identified data from BreastScreen Victoria (BSV), a free mammographic service to Victorian women aged 50-69 years, every two years, with the aim to reduce breast cancer mortality. The program was established in 1992 and currently screens about 160,000 women each year. Participation is approximately 58% of all eligible Victorian women between 50-69 years of age. In this study we used data routinely collected by BSV including information on age, hormone replacement therapy (HRT), family history, symptomatic status, and date of each screen in addition to tumor characteristics: histopathology, tumor size (mm) and histological grade and whether diagnosis was DCIS or invasive cancer.

Data management and coding of relevant variables were conducted using Microsoft ACCESS and STATA. Key variables of interest were time since previous negative screen coded as the time in months from most recent screen (prior to diagnosis) to the previous screen. HRT was coded as HRT use at most recent screen (yes, no) and years of HRT use at most recent screen (None, 1 to 5 years, over 5 years). Age at diagnosis was coded in years. Histological grading is defined as grade 1: Low grade (well differentiated), grade 2: Intermediate grade (moderately differentiated) and grade 3: High grade (poorly differentiated) and unknown grade.

Background work:
i. Review of literature on natural history of DCIS

Prior to commencing the work on simulation and modeling of DCIS, we reviewed the literature on evidence regarding the natural history of DCIS and in particular modeling relevant to the natural history. A manuscript is about to be submitted to the journal Cancer (see reportable outcomes). The available evidence suggests that only a modest proportion of DCIS may progress to invasive cancer; however, all sources of evidence have limitations that may bias the estimates in either direction.

ii. Preliminary analyses characterizing DCIS size and grade in relation to screening interval.
To characterize DCIS size distributions for women attending a subsequent screen we used multiple linear regression methods. The models comprise of an age adjusted base model with each predictor entered in a step-wise manner with exclusion criteria based on the p value for each likelihood ratio test. To investigate associations between HRT use at diagnosis screen and size, the data was stratified by women aged over 55 years. Multinomial regression methods were used to evaluate predictors of histological grade. These methods are extensions of logistic regression where the outcome consists of more than two categories (histological grade consists of three levels; high, intermediate and low grade). Potential predictors of grade were entered into the model in a similar fashion to the predictors of size. All analyses were completed using the statistical package STATA version 7.

In BreastScreen Victoria n = 1127 women were diagnosed with non-invasive breast malignancy – n = 552 Comedo, n = 319 non-comedo, n = 228 mixed and n = 28 other DCIS between 1993 and 2000. Of these, n=590 were diagnosed at the first screen and n=537 were diagnosed at repeat screen. Of the DCIS diagnosis n = 724 women had non-missing tumor size (mm). Table 1 shows results of a multivariate regression of predictors of log-transformed size. Only high grade ($p < 0.001$) was associated with larger lesions. When restricted to subsequent attendees of the program a longer screening interval was not associated with size. Neither HRT use, nor duration of HRT use was associated with DCIS size. Multinomial regression models were constructed to assess the multivariate effects of country of birth, area of residence, symptomatic status, previous benign breast disease, histology, grade, family history, and HRT status on histological grade adjusting for age at diagnosis (Table 2). High grade lesions were 6 times more likely to be larger (> 20mm) in comparison to low grade lesions, $p < 0.001$ (95% CI 2.3 to 15.73). For subsequent attendees of the program time since previous negative screen (months) was not associated with an increased risk of high grade tumors ($p = 0.2$).

These preliminary results suggest that DCIS is relatively slow growing, at least within relatively short screening intervals of several years, as time since screen does not predict size of DCIS. However, an important issue is at what size DCIS is likely to invade and whether this depends solely on size of the tumor.

iii. Simulation and modeling of growth rates of DCIS using BreastScreen data.

The work on simulation and modeling of DCIS to date has progressed well and our aims are to:

a. Simulate size of DCIS at first and subsequent screen under various scenarios of tumour initiation, growth, detection
b. Compare simulated distributions to data from screening program (BSV)
c. Identify scenarios that are compatible or not with BSV data

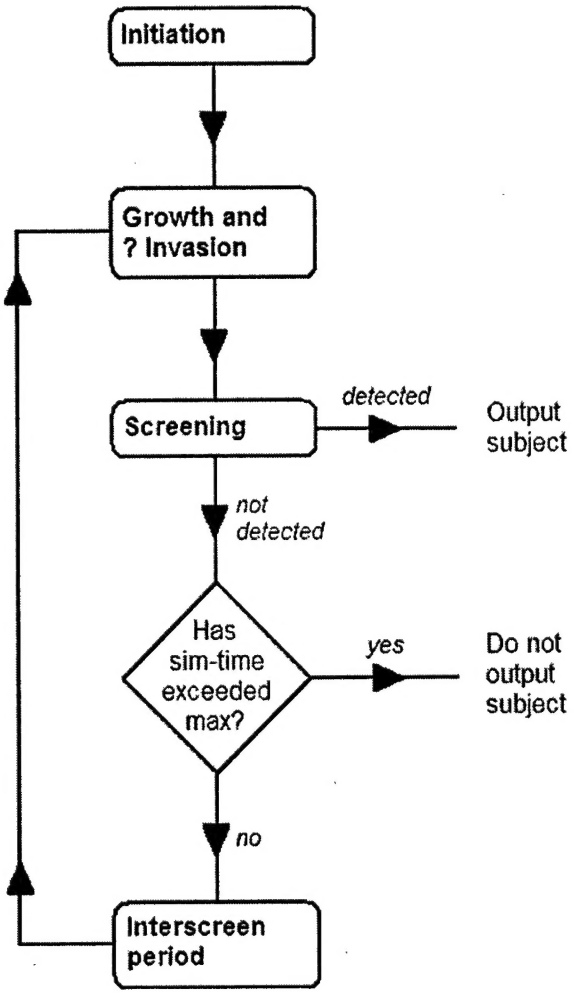
This approach is useful in refuting or ruling out certain hypotheses regarding growth rates of DCIS. We have selected a limited number of different scenarios, for example high growth, high detection and high invasion probabilities vs low growth, low invasion and high detection probabilities. Our aim is to identify those scenarios that best fit the observed data and then ultimately validate this approach using another data set.

The simulation has been implemented in C++ and an executable version is available for the Windows Xp platform. However since the source code is also available the program can be run

on any computer with a C++ compiler. The simulation begins with various menus and prompts asking the user to enter details such as the name of the output file and the number of subjects to simulate. The user is also asked to choose from various options related to tumor initiation, growth and detection. For a description of each of these options, see the section *Configuring the simulation* (below).

The flowchart below (Figure 1) depicts the simulation's control flow for each subject. The program iterates until it has generated the required number of subjects with a detected tumor.

Figure 1



For each subject, the simulation begins by randomly choosing the subject's age at tumorigenesis and at first screen. Age at tumorigenesis is selected from the distribution which Pike *et al*[6] suggest for the incidence of breast cancer, though we estimate key hormonal risk factors such as the distribution of ages at menarche from the control group of the Australian Breast Cancer Family Study[7]. Age at first screen is chosen from a uniform distribution with range 45 to 75 years. This approximates the distribution observed in the BreastScreen Victoria mammography data collected between 1993 and 2000[8].

Next a grade is assigned to the tumor according to a set of probabilities chosen by the user. Growth rates for DCIS and IBC are drawn from grade-specific distributions. If the user has

chosen the logistic growth option then a limiting size is also randomly selected. The size of the tumor is set to a starting value, usually the size of one cell. Each tumor starts as DCIS.

In the first instance, growth and potential invasion are simulated for the period of time from tumor initiation to first screen (unless tumorigenesis occurs after the first screen, as discussed above). After this, growth and potential invasion are simulated for the periods of time between consecutive screenings.

The growth module simulates a period of growth by updating the size of the tumor. Growth is simulated for both DCIS and IBC or just for DCIS, depending on whether invasion has occurred or not. The simulated growth depends on the length of the growth period, the rate of growth, the type of growth (exponential, linear, logistic or power law) and a limiting size if logistic growth has been selected.

We have compared our simulated data to that observed within the screening program. For illustrative purposes we present preliminary results from the simulations and show how these can be used to refute certain hypotheses regarding growth rates of DCIS.

Figures 2 (a) and (b) are quartile-quartile plots comparing the size distributions of simulated tumours (generated under two different scenarios) with the size distribution observed in the mammography data. A good match occurs when all data points lie close to the diagonal line.

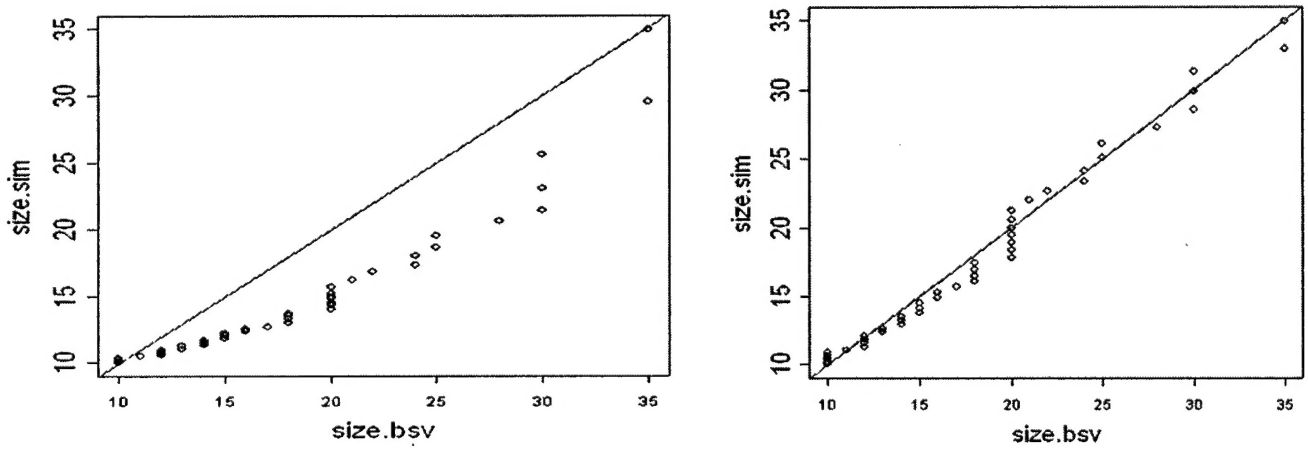


Figure 2 (a) is modeled using assumptions of high detection sensitivity, linear growth, high invasion rate (median size at invasion = 7mm) and high growth rate (median time to reach 10mm = 4 years). It compares simulated data to the distribution of DCIS observed at first screen within the screening program and predicts 1% DCIS detected and mean size of 10.1mm. Figure 2 (b) is modeled using assumptions of high detection sensitivity, linear growth, low invasion rate (median size at invasion = 25mm) and low growth rate (median time to reach 10mm = 10 years). It compares simulated data to the distribution of DCIS size observed at first screen within the screening program and predicts 10% DCIS detected at first screen with mean size = 21.0 mm.

The simulation from Figure 2(b) clearly fits the data better and suggest that a low growth rate, low invasion rate is more consistent with the screening data than high invasion and high growth rates of DCIS. Over the next 12 months we plan to refine these simulations and use distributions of DCIS size by time since screen and first and subsequent screen to test these hypotheses. Once we decide on the scenario and models that best fit the data, we will also use back-calculation techniques to determine the proportion of invasive breast cancers that arise from DCIS.

Task 4: Manuscript preparation

We have completed a review of the literature on natural history of DCIS and this is about to be submitted for publication to *Cancer* , Erbas et al "The natural history of Ductal Carcinoma in situ of the breast: a review".

Several other manuscripts are presently in preparation:
Erbas et al "Trends and predictors of size and grade of DCIS within a screening program"
Dowty et al "A comparison of breast cancer invasion mechanisms".

We will require a further 6-12 months to complete all of the manuscripts that will arise from this project.

Positive and negative findings

Our findings to date are:

1. Existing evidence from the literature suggests that only a modest proportion of DCIS will progress to invasive cancer; however, all data sources are limited and may potentially bias results in either direction.
2. Size and grade of DCIS are not determined by time since previous screen, suggesting that, at least within moderate screening intervals observed in this program, growth rates of DCIS are likely to be slow.
3. Preliminary results from the simulation of observed DCIS size distributions from a screening program suggest that low growth and low invasion rates are most compatible with the observed data.

Problems in accomplishing tasks and recommended changes

Our main difficulties in accomplishing these tasks relate to the short time frame for the project. It has not been possible to complete the validation work and all of the simulation within a 12 month time frame; however substantial progress has been made and we anticipate completion of all tasks, including write-up of manuscripts, within the next 12 months.

Another problem in the modeling work is trying to capture the inherent complexity of the process of growth and invasion. Although our preliminary simulations fit the data quite well, it has proved difficult to adequately simulate the size distribution at subsequent screen. We are exploring the possibility of restricting the models to specific grades of tumors, as it is possible that the simulations may require stratification by tumor grade, which is a known determinant of tumor growth rates. Our results will need to be validated in another data set from a screening program.

KEY RESEARCH ACCOMPLISHMENTS/FINDINGS

1. We have reviewed the literature on natural history of DCIS (manuscript to be submitted shortly). The data suggest that not all DCIS progress to invasive cancer and show that all sources are limited and may bias estimates in either direction.
2. We have analyzed existing data from the screening program to determine predictors of DCIS size and grade. Our results show that grade is the strongest predictor of size and that time since screen does not predict either DCIS size or grade. HRT use is associated with better grade DCIS (similar to invasive cancer) but is not associated with size of DCIS.

3. We have made significant progress in modeling and simulation of observed DCIS size distributions from a screening program under different assumptions of growth, invasion and detection. We show that these simulations provide a useful mechanism to rule out various scenarios DCIS growth and invasion. Preliminary results show that the scenarios assuming low growth and low invasion fit the data better that those assuming a high invasion and high growth rate.

Over the next 12 months we will further refine these simulations incorporating DCIS size distributions at first and subsequent screens and will then use these models to back-calculate the proportion of invasive tumors that

REPORTABLE OUTCOMES

Manuscripts:

Erbas B, Provenzano E, Armes J, Gertig DM The natural history of DCIS of the breast: a review. To be submitted to Cancer 9/04.

Erbas et al Trends and predictors of size and grade of DCIS within a screening program. (in preparation)

Dowty et al A comparison of breast cancer invasion mechanisms. (In preparation)

Byrnes et al Simulation of DCIS detection and invasion: comparison with screening data (In preparation)

Published Conference Proceedings/Abstracts

Erbas B., Chang J-H., Byrnes G., Provenzano E., Kavanagh AK., Gertig DM. Trends and predictors of size and grade for Ductal Carcinoma In Situ diagnosed within a mammographic screening program. (*Proceedings from the Symposium Mammographicum July 2004)*

Erbas B., Chang P., Byrnes G., Provenzano E., Kavanagh AK., Gertig DM. Trends and predictors of size and grade for Ductal Carcinoma In Situ diagnosed within a mammographic screening program. (*Proceedings from the Australasian Epidemiological Association October 2004)*

J. Dowty et al., Investigating the natural history of DCIS using a computer-based simulation (*Proceedings from the Australasian Epidemiological Association October 2004)*

Awards

"Natural history of Ductal carcinoma in situ of the breast: using statistical models to estimate growth and progression". Faculty of Medicine, Dentistry & Health Sciences Annie S Glover Research Fellow Award in Cancer (PI: Bircan Erbas) $70,000 + on costs.

Pending

"Novel applications of statistical methods to mammographic screening data". Department of Defense-Multidisciplinary postdoctoral award (PI: Bircan Erbas) $401,170US requested for 2005-2007

"Novel Applications of Statistical Methods to Breast Cancer Data". National Health Medical Research Council-postdoctoral fellowship (PI: Bircan Erbas) Full time salary requested for 2005-2008

LIST OF PERSONNEL RECEIVING PAY

Dr Bircan Erbas (Post-doctoral fellow)

Ms Annette Fedson (Research assistant)

Ms Colleen Code (Statistician)

Dr Dorota Gertig (Senior Research Fellow)

CONCLUSIONS

New approaches to estimate the natural history of DCIS are essential. The aim of this study is to use novel applications of statistical methods to estimate the proportion of DCIS that progress to invasive cancer. We have developed a computer simulation for mammographic screening data which models progression and detection of Ductal carcinoma in situ. Based on various options for growth, detection and invasion, we have simulated various distributions of DCIS sizes for a screening program. These distributions can then be used to test hypotheses regarding different scenarios of growth and invasion of DCIS. Preliminary results show that low growth rates and low invasion rates provide the best fit to the data. Further work will include the addition of screening round and different mechanisms of invasion to the modeling.

REFERENCES

1. Betsill WL, Rosen PP, Lieberman PH et al. Intraductal carcinoma. Long term followup after treatment by biopsy alone. JAMA 1978;239:1863-7
2. Eusebi V, Feudale E, Foschini MP. Long-term follow-up of in-situ carcinoma of the breast. Semin Diagn pathol 1994;11;223-35
3. Page DL, Dupont WD, Rogers LW, Landenberger M. Intraductal carcinoma of the breast after follow-up by biopsy alone 1982;49:751-8
4. Hart D, Schochat E, Agur Z. The growth law of primary breast cancer as inferred from mammographic screening trials data. Br J Cancer 1998;78:382-7
5. Brown BW, Atkinson EN, Bartoszynski R et al. Estimation of human tumor growth rate from distribution of tumour size at diction. J Natl Cancer Inst 1984;72:31-8
6. Pike MC, Krailo MD, Henderson BE et al "Hormonal risk factors, breast tissue age and the age-incidence of breast cancer'. Nature. 1983 Jun 30;303(5920):767-70.
7. Hopper JL, Chenevix-Trench G, Jolley DJ, Dite GS, Jenkins MA, Venter DJ, McCredie MR, Giles GG.Design and analysis issues in a population-based, case-control-family study of the genetic epidemiology of breast cancer and the Co-operative Family Registry for Breast Cancer Studies (CFRBCS).J Natl Cancer Inst Monogr. 1999;(26):95-100.
8. BreastScreen Victoria Decade Statistical Report. 2003.

APPENDIX 1: Predictors of size and grade of DCIS. Results from preliminary analyses of existing BreastScreen data.

Table 1: Results of multivariate regressions analysis of log size and HRT use and various risk factors adjusting for age at diagnosis for screen-detected DCIS within BreastScreen Victoria

	Variables	Beta coeff. Estimate #	s.e. #	95% CI #	P value#
	Age				
	50-69 yrs	--	--	--	--
	< 50 yrs	0.29	0.15	(-0.01, 0.60)	0.06
	> 69 yrs	-0.07	0.13	(-0.32, 0.18)	0.59
	Country of birth				
	Australasian	--	--	--	--
	UK	0.12	0.16	(-0.18, 0.43)	0.43
	Europe	0.13	0.13	(-0.13, 0.38)	0.32
	Asian	-0.09	0.20	(-0.49, 0.31)	0.66
	Other	-0.03	0.24	(-0.49, 0.43)	0.90
	Symptoms				
	No	--	--	--	--
	Ever	-0.20	0.18	(-0.55, 0.15)	0.27
	Postcode				
	Capital city/ major urban	--	--	--	--
	Rural/remote	-0.11	0.13	(-0.36, 0.14)	0.39
	Previous breast disease				
	No	--	--	--	--
	Ever	0.02	0.10	(-0.18, 0.22)	0.83
	Histology				
	Comedo	--	--	--	--
	Non-comedo	-0.14	0.15	(-0.43, 0.16)	0.37
	Mixed	0.07	0.13	(-0.17, 0.32)	0.56
	Other	0.03	0.27	(-0.50, 0.57)	0.90
	Grade				
	Low	--	--	--	--
	Medium	0.43	0.15	(0.13, 0.73)	0.005
	High	0.75	0.17	(0.42, 1.08)	<0.001
	Family history				
	First degree	--	--	--	--
	Second degree	0.04	0.22	(-0.40, 0.48)	0.85
	Unknown/other	0.21	0.20	(-0.19, 0.61)	0.30
	Time since last negative screen **	0.01	0.01	(-0.02, 0.04)	0.42
Age >=55yrs	**HRT use**				
	No	--	--	--	--
	Ever	0.19	0.14	(-0.09, 0.46)	0.18
	HRT duration				
	None	--	--	--	--
	1-5 yrs	0.24	0.23	(-0.21, 0.70)	0.30
	> 5 yrs	0.16	0.16	(-0.15, 0.48)	0.31
	Time since last negative screen **	0.02	0.02	(-0.02, 0.06)	0.41

Table 2: Results of multinomial regression of histological grade and various risk factors adjusting for age at diagnosis for screen-detected DCIS within BreastScreen Victoria

		Grade 3 vs 1		Grade 2 vs 1	
	Variables	**OR (95% CI) #**	**P value**	**OR (95% CI) #**	**P value**
ALL WOMEN:					
	Age				
	50-69 yrs	--	--	--	--
	< 50 yrs	0.83 (0.25, 2.72)	0.76	0.74 (0.26, 2.10)	0.57
	> 69 yrs	1.01 (0.39, 2.62)	0.99	1.17 (0.50, 2.73)	0.71
	Country of birth				
	Australasian	--	--	--	--
	UK	1.02 (0.29, 3.56)	0.97	0.61 (0.19, 1.94)	0.40
	Europe	1.27 (0.45, 3.54)	0.65	1.09 (0.42, 2.80)	0.86
	Asian	1.64 (0.33, 8.12)	0.54	1.53 (0.39, 6.01)	0.54
	Other	2.63 (0.32, 21.80)	0.37	2.25 (0.33, 15.38)	0.41
	Symptoms				
	No	--	--	--	--
	Ever	1.62 (0.40, 6.55)	0.50	1.56 (0.44, 5.49)	0.49
	Postcode				
	Capital city/ major urban	--	--	--	--
	Rural/remote	1.64 (0.57, 4.74)	0.36	1.1 (0.40, 3.09)	0.84
	Previous breast disease				
	No	--	--	--	--
	Ever	0.70 (0.33, 1.52)	0.37	0.68 (0.34, 1.38)	0.29
	Histology				
	Comedo	--	--	--	--
	Non-comedo	0.004 (0.001, 0.01)	<0.001	0.09 (0.03, 0.23)	<0.001
	Mixed	0.29 (0.08, 1.01)	0.05	0.82 (0.22, 2.98)	0.176
	Other	0.03 (0.007, 0.14)	<0.001	0.10 (0.02, 0.48)	0.004
	Size				
	< 10 mm	--	--	--	--
	10-20 mm	2.14 (0.90, 5.06)	0.08	1.20 (0.53, 2.71)	0.66
	> 20 mm	6.02 (2.30, 15.73)	<0.001	3.34 (1.38, 8.08)	0.007
	Family history				
	First degree	--	--	--	--
	Second degree	1.09 (0.21, 5.58)	0.92	2.52 (0.54, 11.82)	0.24
	Unknown/other	1.55 (0.35, 6.90)	0.57	2.11 (0.50, 8.86)	0.31
	Time since last negative screen **	0.98 (0.87, 1.10)	0.73	0.96 (0.85, 1.08)	0.49
Age ≥ 55:					
	HRT use				
	No	--	--	--	--
	Ever	0.48 (0.16, 1.44)	0.19	0.51 (0.18, 1.39)	0.19
	HRT duration				
	None	--	--	--	--
	1-5 yrs	0.40 (0.08, 2.02)	0.27	0.54 (0.14, 2.14)	0.38
	> 5 yrs	0.51 (0.14, 1.91)	0.32	0.48 (0.14, 1.69)	0.26
	Time since last negative screen **	0.87 (0.73, 1.05)	0.14	0.90 (0.76, 1.06)	0.21